

Halldóra Jónsdóttir og Þórdís Úlfarsdóttir

Íslensk nútímamálsorðabók. Kjarni tungumálsins

„[...] tilgangur [orðabókarinnar] er að veita ófalsaða og sanna mynd af tungumálinu eins og það er.“

(Sigfús Blöndal 1920–1924:xi, þýð. höf.)

1 Inngangur

Undanfarin ár hefur verið unnið að gerð nýrrar íslenskrar orðabókar hjá Stofnun Árna Magnússonar í íslenskum fræðum (SÁM) sem ber heitið *Íslensk nútímamálsorðabók*. Ekki er um að ræða prentaða bók heldur rafrænt verk sem frá upphafi er hannað fyrir birtingu á vefnum. Orðabókin er meðalstór og hefur að geyma um 50.000 uppflettiorð. Hún er að mestu byggð á annarri veforðabók sem einnig er gefin út hjá SÁM, margmála skandinavísku orðabókinni ISLEX sem orðin er vel þekkt meðal notenda sinna. Þessi grein fjallar um *Íslenska nútímamálsorðabók*, upphaf hennar, helstu nýmæli og ritstjórnarstefnu. Leitast verður við að svara spurningum sem varða tilurð og tilgang orðabókarinnar og þær nýjungar sem þar er að finna, og jafnframt hvernig vinnubrögð hafa breyst miðað við orðabækur frá því fyrir aldamótin 2000. Meðal nýrra aðferða er að vinna orðaforðann eftir merkingarflokkum en ekki í stafrófsröð, svo og að nýta fjölbreytileg textasöfn sem grundvallartæki við orðabókarvinnuna. Enn fremur er orðabókin frá upphafi hönnuð fyrir vefinn sem býður upp á hagnýtingu á ýmiss konar margmiðlunarefni.

Sumir kunna að spyrja hvort yfirleitt hafi verið þörf á nýrri íslenskri

orðabók og hvort sú orðabók sem lengi hefur verið til, *Íslensk orðabók* (ÍO) (frumútgáfa 1963, 2. útgáfa 1983, 3. útgáfa 2002 og 4. útgáfa 2007; sjá nánar formála ÍO 2007), ætti ekki að duga flestum sem þurfa á upplýsingum um tungumálið að halda. Sú orðabók hefur þjónað Íslendingum í yfir fimmtíu ár, í mörgum útgáfum, bæði prentuðum og í rafrænum aðgangi frá 2000. En eins og Sigfús Blöndal (1924:vii) segir í formála sínum að *Íslensk-danskri orðabók* „hefur menningarleg framþróun kallað á mikinn fjölda nýrra hugtaka og orða á flestum sviðum sem annað hvort eru nýyrði eða erlend tökuorð“, og á það mjög við á okkar tímum rétt eins og fyrir um hundrað árum. Íslenska er lifandi tungumál sem tekur stöðugum breytingum, þótt menn verði ekki alltaf varir við það, og það kallar á nýjar orðabækur, rétt eins og önnur fræðirit. Slík rit hafa þó þann eiginleika að úreldast fremur hratt eftir því sem breytingar verða á hinum ýmsu sviðum samfélagsins, og það er ekki fyrr en eftir aldamótin 2000 sem nokkur leið er að halda í við þróunina, m.a. með tilkomu veforðabóka þar sem hægt er að sinna stöðugum uppfærslum. Jafnframt er vafasamt að tala um að slíku verki sé yfirleitt lokið ef ætlunin er að orðabókin elti uppi tungumálið og haldist nokkurn veginn í hendur við málbreytingar og nýmyndanir.

Efni greinarinnar skiptist þannig: Í kafla 2 er rakin saga íslenskra orðabóka í mjög stórum dráttum. Kafli 3 lýsir efni *Íslenskrar nútímamálsorðabókar* og er þar greint frá helstu innviðum hennar og eiginleikum. Kafli 4 fjallar um ritstjórnarstefnu eins og hún birtist í vali orðaforðans og orðskýringum. Þar segir einnig frá notkun á orðabókinni og uppfærslum á efni hennar. Að endingu eru niðurstöður dregnar saman í 5. kafla.

2 Eldri orðabækur

Hér verður í stuttu máli sagt frá fleiri þáttum í sögu íslenskrar orðabókaútgáfu, en upphaf hennar má rekja til 17. aldar.

2.1 Fyrir 20. öldina

Orðabækur spretta ekki upp án samhengis heldur byggja þær á menningarlegum og sögulegum grunni hvers tíma og fyrstu íslensku orðabækurnar áttu sér erlendar fyrirmyndir. Hér á landi sem annars staðar tók nokkrar aldir að móta framsetningu orðabóka í þá átt sem

þekkest í dag. Á fyrri stigum var t.d. stafrófsröðin ekki alltaf hin sjálf-gilda röð uppflöttiorðanna. Í *Nucleus latinitatis* Jóns Árnasonar (1738) er latnesku orðunum raðað eftir orðstofnum, þ.e. skyld orð standa saman undir einni höfuðflettu. Sagnorð voru ekki færð upp í nafnhætti (*sitja, fá*) heldur í 1. persónu eintölu (*sit, fæ*) samkvæmt hefð í latneskum orðabókum erlendis. Það er einnig svo í orðabók Björns Halldórssonar (1992 [1814]) þótt uppflöttiorðin í þeirri bók séu íslensk. Stafrófsröð bókstafanna var nokkuð öðruvísi en nú tíðkast, og er t.d. stöfunum *o, ó* og *ö* blandað saman í síðarnefnda ritinu.

Fyrir 19. öldina er íslensk orðabókasaga ekki mikil að umfangi. Elstu orðabækurnar eru ekki einsmálsbækur heldur hafa þær annað tungumál en íslensku annað hvort sem grunnmál eða markmál, oftast latínu eða dönsku. Allra elsta orðabókin, *Lexicon Islandicum* frá 1683 (Guðmundur Andrússon 1999 [1683]) er með íslensk uppflöttiorð og skýringar á latínu, og skólaorðabókin *Nucleus latinitatis* frá 1738 hefur uppflöttiorð á latínu og skýringar á íslensku (Jón Árnason 1994 [1738]). Á tímabilinu 1700–1900 komu að öðru leyti einkum út orðabækur yfir forníslensku og fornt skáldamál, með skýringum á dönsku, norsku eða ensku.

Á 19. öld virðist hafa orðið mikil vakning í orðabókagerð á Íslandi þrátt fyrir að íslenskir málnotendur væru mjög fáir, en um miðja öldina voru Íslendingar aðeins um sextíu þúsund.¹ Í ljósi þessa fámennis er eftirtektarvert að á þeim tíma komu út allmargar orðabækur. Var það án efa vegna áhrifa frá öðrum löndum Evrópu þar sem orðabókagerð var orðin útbreidd iðja. Á þessum tíma var Ísland ekki fullvalda ríki heldur hjálenda Danmerkur, en þjóðernisvakning landsmanna og virðing fyrir tungumálinu kann að hafa haft hvetjandi áhrif til útgáfu orðabóka þar sem íslenskt mál var viðfangsefnið. Bækurnar voru venjulega prentaðar og gefnar út erlendis. Það skal tekið fram að hér er átt við hefðbundnar orðabækur með orðskýringum, eða þýðingum ef um tvímálarit var að ræða, en ekki rit um stafsetningu, orðsifjar o.fl.

Nefna má þrjú dæmi um stórar og vandaðar orðabækur sem komu út á 19. öld. Í fyrsta lagi *Íslensk-latnesk-dönsk orðabók* eftir Björn Halldórsson (1992 [1814]), sem hann nefndi latneska heitinu *Lexicon Islandico - Latino - Danicum*. Í öðru lagi orðabók Konráðs Gíslasonar, *Dönsk orðabók með íslenzkum þýðingum* sem kom út 1851, og í þriðja lagi kom út *Ný dönsk orðabók með íslenzkum þýðingum* eftir Jónas Jónasson

¹ Sbr. vef Hagstofu Íslands.

árið 1896. Nokkrar fleiri orðabækur urðu til á 19. öld sem ekki verður fjallað um hér. En gerð tvímálaorðabóka hélt áfram af krafti fram yfir aldamótin 1900 og stöðugt eftir það.

2.2 Orðabækur á 20. öld

Á 20. öld voru gefnar út allmargar orðabækur milli íslensku og annarra mála, einkum þeirra tungumála sem lengst hafa verið kennd í skólum á Íslandi, m.a. dönsku, ensku, þýsku og frönsku. Þessar orðabækur eru mjög misjafnar að stærð og gæðum og verður ekki farið nánar út í þá sálma hér. Þó verður að nefna sérstaklega *Íslensk-danska orðabók* (1920–1924) sem kom út snemma á 20. öld og sagt er frá í næsta kafla. Hún markar upphaf nútímalegrar orðabókagerðar þar sem íslenska er í brennidepli.

2.2.1 *Íslensk-dönsk orðabók* eftir Sigfús Blöndal

Í upphafi 20. aldar var hafið blómaskeið stórra, sögulegra orðabóka í mörgum Evrópulöndum. Upp úr aldamótunum 1900 hófst í Danmörku útgáfa á *Ordbog over det danske sprog*, en fyrsta bindi þess verks (af 28 bindum alls) kom út 1919. Í Bretlandi var þegar hafin útgáfa á *Oxford English Dictionary* (1884), og fyrsta bindi *Svenska Akademiens Ordbok* kom út árið 1898 í Lundi.² Sama máli gegndi um stórar þýskar og hollenskar orðabækur.

Hin efnismikla orðabók Sigfúsar Blöndals, *Íslensk-dönsk orðabók* (1920–1924), er fyrsta verkið þar sem leitast er við að setja saman „þjóðarorðabók“ fyrir íslenskt mál, þ.e. yfirgripsmikið verk sem nær yfir mikinn orðaforða, hefur að geyma notkunardæmi og orðasambönd, og er ítarleg lýsing á tungumálinu eins og það var talað og skrifað á ritunartíma verksins, þ.e. fyrstu áratugum 20. aldar. Þannig var því tíðarandinn þegar Sigfús Blöndal vann að orðabókinni ásamt eiginkonu sinni Björgu Þorláksdóttur Blöndal, Jóni Ófeigsyni, Holger Wiehe og fleiri samstarfsmönnum. Sigfúsi Blöndal var umhugað um að orðabók hans endurspeglaði orðaforða og málnotkun samtímans, og í formálanum segir hann: „[...] þrátt fyrir að einhverjar eldri orðabækur hafi að vissu leyti tekið tillit til nútímamálsins, þá hefur skortur á orðabók, sem gæti varpað ljósi á allar hliðar tungumálsins, sífellt orðið áþreifanlegri. [...] Menn fundu

² Því verki er raunar enn ekki lokið en árið 2018 hafa verið gefin út 37 bindi sem ná yfir stafkaflana A–V.

ekki einungis fyrir þessum skorti innan fræðasamfélagsins, heldur einnig við dagleg störf“ (Sigfús Blöndal 1920–1924:vii). Í orðabókargreinunum er að finna fjöldann allan af mállýskuorðum, orðasamböndum og notkunardæmum, miklu meira en nokkurn tíma hafði áður sést í íslenskum orðabókum. Sérstök áhersla var lögð á íslenskt alþýðumál. Bókin átti ekki að vera orðabók fyrir fræðimenn heldur átti hún að lýsa íslensku máli eins og það var talað og ritað. Orðabókin var fyrst og fremst samin fyrir Íslendinga en af sögulegum og útgáfulegum ástæðum voru orðskýringarnar hafðar á dönsku. Enda þótt orðskýringarnar væru vandaðar hljóta þær að hafa dregið nokkuð úr notagildi orðabókarinnar fyrir allan almenning sem kunnir dönsku misjafnlega vel.

Íslensk-dönsk orðabók er afar efnismikil, yfir þúsund blaðsíður, og má segja að með henni hafi orðið þáttaskil í íslenskri orðabókagerð, bæði hvað varðar fagmennsku í vinnubrögðum og umfang ritsins. Sigfús hafði hugsað sér að orðabókin yrði endurskoðuð og endurútfegin reglulega, og átti sérstakur sjóður sem stofnaður var að fjármagna það verk. En eins og kunnugt er var það aldrei gert og aðeins ein útgáfa af *Íslensk-dönskri orðabók* hefur litið dagsins ljós.³

2.2.2 *Íslensk orðabók handa skólum og almenningi*

Fyrsta íslenska einsmálsorðabókin, þ.e. orðabók þar sem bæði uppflettiordin og skýringarnar eru á íslensku, kom ekki út fyrr en eftir miðja 20. öld. Það var *Íslensk orðabók handa skólum og almenningi* í ritstjórn Árna Böðvarssonar sem var fyrst gefin út hjá Menningarsjóði árið 1963. Önnur útgáfa hennar, talsvert aukin og bætt, kom út tuttugu árum síðar. Menningarsjóður var ríkisrekið forlag sem gaf út bæði fræðirit og fagurbókmenntir. Árið 1992 var Menningarsjóður lagður niður og var þá útgáfuréttur að *Íslenskri orðabók* boðinn út, og í kjölfarið seldur til bókaforlags Máls og menningar. Árið 1996 var hafist handa við að koma *Íslenskri orðabók* á gagnagrunnstækt form, og þar með var rafvæðing orðabókarinnar hafin. Hún var að hluta til endurskoðuð næstu árin á eftir, í ritstjórn Marðar Árnasonar, og gefin út á geisladiski árið 2000. Síðan þá hafa komið út nokkrar útgáfur af orðabókinni, bæði sem prentaðar bækur og vefútgáfur. Þess má geta að með tilkomu gagnagrunna var hætt að blanda saman breiðum og

³ Þó var gefinn út viðbætur við orðabókina árið 1963 í ritstjórn Halldórs Halldórssonar og Jakobs Benediktssonar, og voru þar viðbótarorð sem ekki voru í upphaflegu bókinni. Enn voru allar orðskýringarnar á dönsku.

grönnum sérhljóða í stafrófsröð uppflettiorðanna (t.d. *a* og *á*).⁴ Við tölvuvinnslu orðabóka var slík blöndun bókstafa ekki lengur sjálfgild eða hentug, og er ólíklegt að nokkur sakni þeirrar röðunar nú.

Í yfir fimm áratugi var *Íslensk orðabók* eina íslenska einsmálsorðabókin sem vöð var á fyrir notendur. Greinarhöfundar störfuðu báðir við endurskoðun *Íslenskrar orðabókar* í nokkur ár eftir að hún var flutt í tölvutækt form, og reyndist það dýrmæt reynsla þegar ráðist var í að búa til önnur orðabókaverk, einkum ISLEX og *Íslenska nútímamálsorðabók*.

2.2.3 ISLEX

Veforðabókin ISLEX⁵ er sá grundvöllur sem *Íslensk nútímamálsorðabók* er byggð á. ISLEX er margmála orðabók með íslenskum uppflettiorðum og sex norðurlandamálum sem markmálum: dönsku, norsku (bæði nýnorsku og bókmáli), sænsku, færeysku og finnsku. ISLEX er fyrsta íslenska orðabókin sem unnin er eingöngu fyrir vefinn. Það er samstarfsverkefni sex fræðastofnana: á Íslandi, í Danmörku, Noregi, Svíþjóð, Færeyjum og Finnlandi. ISLEX var opnuð á vefnum árið 2011, en þetta er fyrsta íslenska orðabókin sem eingöngu er ætluð til rafrænnar birtingar. Einnig er hún eina núverandi orðabókin sem nær til svo margra tungumála. Nánari lýsingu á ISLEX-verkefninu er að finna hjá Þórdísi Úlfarsdóttur (2013).

Efni orðabókarinnar, m.a. uppflettiorð, orðskýringar, orðasambönd og notkunardæmi, hentaði mjög vel til að nota í fleiri tvímála-orðabækur, og varð það grunnurinn að nýrri íslensk-franskri orðabók sem hefur verið í vinnslu frá 2014 í samstarfi SÁM og Stofnunar Vigdísar Finnbogadóttir í erlendum tungumálum. Íslenska ritstjórnin hafði enn fremur fært inn í ISLEX orðskýringar til að samræma og auðvelda starf þýðenda í hverju landi fyrir sig, sem varð grunnur að íslenska skýringarorðaforðanum í einsmálsorðabókinni.

3 *Íslensk nútímamálsorðabók*

Byrjað var að vinna við *Íslenska nútímamálsorðabók* snemma árs 2014. Þetta er fyrsta íslenska einsmálsorðabókin í yfir fimmtíu ár sem er

⁴ Í *Íslenskri samheitaorðabók* frá 1985 er þó gerður greinarmunur á grönnum og breiðum sérhljóðum.

⁵ Sjá www.islex.is.

búin til frá grunni en ekki byggð á fyrri útgáfu, og enn fremur er þetta aðeins önnur alíslenska orðabókin frá upphafi vega. Tilurð hennar má rekja til ISLEX-verkefnisins eins og áður segir, en þegar ISLEX var tilbúið var ljóst að safnast hafði mikið og dýrmætt efni sem fyrir utan tvímálaorðabækur var einnig hentugt að nota sem undirstöðu fyrir lýsingu á íslensku máli þar sem markhópurinn væri Íslendingar og aðrir sem nota íslensku. Í hinni nýju einsmálsorðabók urðu því valdir þættir úr ISLEX meginuppistaðan.

Verkinu var í upphafi komið fyrir í eigin gagnagrunni. Fyrirliggjandi efniviður var mikill: flettiorðalisti með fimmtíu þúsund orðum, orðastæður, orðasambönd, notkunardæmi, merki um orðhlutaskil, tenglar í beygingar, og enn fremur listar með samsetningum grunnorða. Upplýsingar um fallstjórn sagna og forsetninga voru tilbúnar ásamt millivísunum innan orðabókarinnar. Myndir voru víða í orðabókargreinunum og framburður uppflöttiorðanna var í formi hljóðskráa. Einum mikilvægasta og tímafrekasta þættinum var jafnframt að mestu lokið, sem fólst í því að skipta uppflöttiorðunum í merkingarliði og búa til skýringar við þau, eina eða fleiri eftir fjölda liðanna. Orðskýringarnar voru þó ekki nógu vel unnar enda höfðu þær upphaflega ekki verið ætlaðar til birtingar.

3.1 Efni orðabókarinnar

3.1.1 Orðaforðinn

Orðaforðinn í *Íslenskri nútímamálsorðabók* á sér nokkra forsögu. Árið 1994 réðst Orðabók Háskólans í samvinnu við Norræna málstöð í Osló í að semja íslenskan orðabókarstofn sem orðið gæti grunnur að tvímálaorðabókum milli íslensku og annarra norðurlandamála. Til verksins fékkst styrkur til tveggja ára frá Norræna menningarsjóðnum. Orðabókarstofninn skyldi vera um fimmtíu til sjötíu þúsund orð. Í stofninum var listi með völdum orðum með áherslu á nútímamálið, en hverju orði fylgdu upplýsingar um orðflokk, formgerð (grunnorð eða samsett orð) auk heimildar um orðið. Kristín Bjarnadóttir hafði umsjón með þessu verki (sjá Kristínu Bjarnadóttur 1998). Þetta varð grunnorðaforði ISLEX-orðabókarinnar.

Þegar vinna við *Íslenska nútímamálsorðabók* hófst var notaður sami listi uppflöttiorða og í ISLEX, en þau voru um fimmtíu þúsund. Á árunum sem liðin voru hafði flettulistinn breyst töluvert frá upphaflegri gerð því að þegar listinn var tekinn saman voru tíðniupplýsingar

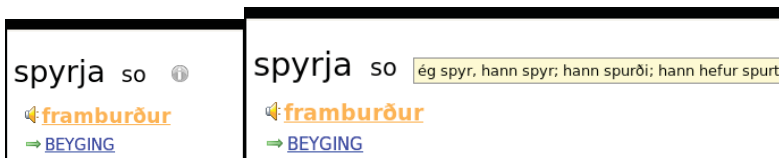
um orðin af skornum skammti. Þurfti því stöðugt að endurmeta orðaforðann eftir því sem verkinu vatt fram. Sum orðin voru talin óþörf og var sleppt en öðrum orðum var bætt við. Af upphaflega orðabókarstofninum hafa 7.900 orð verið fjarlægð en álíka mörg hafa bæst við í staðinn. Það jafngildir um 16% heildarorðaforðans.

Nýi orðaforðinn hefur komið inn í orðabókina aðallega eftir tveimur leiðum: 1) úr umhverfinu og daglegri umræðu, og 2) með orðtöku í þeim tilfellum þar sem vitað var um talsverðar eyður í orðaforðanum, t.d. í greinum eins og jarðfræði og læknisfræði. Enn fremur þurfti að þetta orðaforðann í landaheitum, þjóðarheitum, skólafögum og á fleiri sviðum. Meira verður fjallað um val orðaforðans í tengslum við almenna ritstjórnarstefnu orðabókarinnar í kafla 4.

3.1.2 Beygingar

Beygingar uppflettiörða eru gefnar með tenglum í *Beygingarlýsingu íslensks nútímamáls* (BÍN), sem er gagnagrunnur með beygingum 285.000 íslenskra orða. Þetta fyrirkomulag kemur í stað hinnar hefðbundnu leiðar í prentuðum orðabókum sem er sú að sýna beygingarendingar eða beygingarflokk fyrir aftan uppflettiörðin. Þegar smellt er á tengilinn „beyging“ opnast beygingardæmi viðkomandi orðs í heild sinni. Tenglar eru við öll beygjanleg uppflettiörð: nafnorð, lýsingarorð, sagnir og fornöfn, þ.e. öll orðin í orðabókinni nema smáorð eins og forsetningar, samtengingar og upphrópanir.

Fyrir utan þetta er hægt að sjá ákveðnar orðmyndir allra sagna og lýsingarorða með því að renna mús yfir lítið tákn með bókstafnum „i“ hægra megin við uppflettiörðið. *Mynd 1* sýnir beygingu sagnorðsins *spyrja* en í undirbúningi er að sýna kennimyndir nafnorða á sama hátt.



Mynd 1. Vinstra megin er sögnin *spyrja*. Hægra megin er búið að renna mús yfir i-táknið við hlið flettunnar og birtast þá valdar orðmyndir sagnarinnar.

Orðaleitin í notendaviðmóti orðabókarinnar er jafnframt tengd við gagnagrunn BÍN þannig að ekki þarf að slá inn grunnmynd orðsins. Þetta er mikilvægt fyrir notendur og nýtist t.d. sérstaklega vel nemendum sem læra íslensku sem annað mál sem og þeim sem ekki eru vel að sér í málfræði.

3.1.3 Framburður

Framburður uppflöttiorðanna er einn þeirra þátta sem fenginn er frá ISLEX. Hægt er að hlusta á hljóðupptökurnar með því að smella á hátalaratáknið fyrir neðan viðkomandi uppflöttiorð. Upptökur fóru fram haustið 2012 í hljóðveri þar sem orðin voru lesin upp og hljóðrituð. Auk stakra uppflöttiorða voru lesin upp 752 orðasambönd sem koma fyrir undir kjarnaorðum sínum í orðabókargreinunum.⁶

3.1.4 Orðastæður og orðasambönd

Í *Íslenskri nútímamálsorðabók* er gerður greinarmunur á tvenns konar fastri orðanotkun, annars vegar *orðastæðum* og hins vegar *orðasamböndum*. Á öðrum vettvangi er hvort tveggja nefnt orðasamband en við gerð rafrænu orðabókanna hefur hentað betur að skilja þetta tvennt að (sjá nánar um skilgreiningu orðanna í *Nordisk leksikografisk ordbok* 1997).

Orðastæða er íslenska heitið á því sem á ensku nefnist *collocation*. Það er fast samband tveggja eða fleiri orða þar sem eitt orðið (eða fleiri) er kjarnaliður orðastæðunnar. Merkingin er almennt gegnsæ: *rjóður í kinnum, kulnað eldfjall, kinka kolli, ganga í klaustur*. Orðasamband nefnist á ensku *fixed expression* (eða *idiom*). Merking orðasambandsins getur verið alveg ógegnæ þar sem kjarni þess hefur oft lítið eða ekkert að gera með merkinguna í heild: *vaxa úr grasi, láta greipar sópa*. Hér er um að ræða myndhverf sambönd og stundum hrein orðtök. Ítarlegri fróðleik um mismunandi tegundir orðasambanda er að finna hjá Jóni Hilmari Jónssyni (2005:xix–xxi).

Orðasambönd og orðastæður hafa erfst í gegnum ISLEX, en langmestur hluti þeirra á upptök sín í *Stóru orðabókinni um íslenska málnotkun* (Jón Hilmar Jónsson 2005). Mörg þessara sambanda eru þó ekki nákvæmlega eins og í upprunalega verkinu heldur hafa þau víða verið einfölduð, enda eru ólíkar orðabækur hér á ferð. Alls eru orðastæður og orðasambönd í *Íslenskri nútímamálsorðabók* um 10.000 talsins.

⁶ Orðin voru lesin upp af spjaldtölvu til hagræðis og var því ekkert skrif í blöðum sem truflaði né heldur þurfti að fletta á milli síðna. Lesari var Guðfinna Rúnarsdóttir og upptökur fóru fram í hljóðverinu Upptekið ehf.

3.1.5 Notkunardæmi

Mörg notkunardæmi eru í orðabókinni. Þeim er ætlað að sýna uppflettiorðið í eðlilegu samhengi og einnig að skerpa greinarmuninn á mismunandi merkingum orðsins, sé um fleiri en einn merkingarlið að ræða. Helsta uppspretta dæmanna eru textasöfn (sjá kafla 3.2.2) en þau hafa nýst mjög mikið til að búa til dæmin. Það hefur þó ekki tíðkast að taka heilar setningar úr textum og gera þær óbreyttar að notkunardæmum heldur er þeim breytt eftir hentugleikum og oft eru gerðar stuttar setningar úr löngum. Ítarlegri greinargerð um notkunardæmin er að finna hjá Þórdísi Úlfarsdóttur (2013:56–57).

3.1.6 Myndir

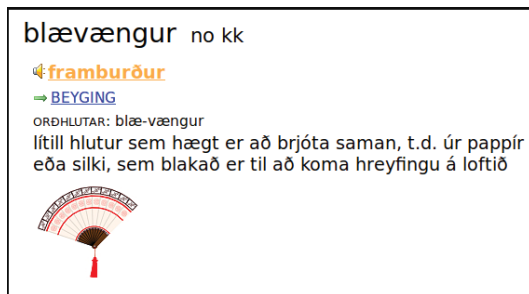
Íslensk nútímamálsorðabók er myndskreytt og má þar finna 3.200 myndir. Mynd getur skýrt merkingu sumra orða betur en hægt er að gera með orðum (*blævængur, íkorni, skæri, tannhjól, tvinnakefli, þvottaklemma*), og stundum er hentugt að hafa myndir til að greina í sundur mismunandi merkingar tiltekins orðs, séu þær fleiri en ein (sbr. t.d. orðið *keila* sem getur verið fisktegund, umferðarkeila og fleira).

Myndirnar koma úr ýmsum áttum en nefna má að sjö hundruð myndir eru eftir Jón Baldur Hlíðberg af fyrirbærum úr ríki náttúrunnar: plöntum, fuglum, fiskum og fleiri dýrum eins og sjá má á *Mynd 2*.



Mynd 2. Myndskreytingar af dýrum og plöntum eftir Jón Baldur Hlíðberg.

Einnig eru í orðabókinni ljósmyndir og teiknaðar myndir. Þær hafa verið fengnar hjá myndabönkum eins og t.d. á *Mynd 3* eða teknar sérstaklega fyrir rafrænu orðabækurnar.



Mynd 3. Skýring við orðið *blævængur* ásamt mynd.

Myndir hafa vitanlega þekkt lengi í orðabókum. Þær geta gert orðabókina meira aðlaðandi í augum notenda, einkum þeirra sem yngri eru.

3.2 Vinnulag: merkingarflokkar og textasöfn

3.2.1 Merkingarflokkar

Íslensk nútímamálsorðabók var unnin eftir merkingarflokkunum en ekki í stafrófsröð eins og áður tíðkaðist. Öllum orðaforða orðabókarinnar er því skipt í merkingarflokka og sér gagnagrunnurinn um að halda utan um þá. *Tafla 1* sýnir dæmi um merkingarflokka.

Að miklu leyti nafnorð	Að miklu leyti lýsingarorð	Að miklu leyti sagnorð
STARFSHEITI	ALDUR	BREYTING
HÚSBÚNAÐUR	ÁNÆGJA	DAUÐI
DÝR	ÁSTAND	DVÖL
JARÐFRÆÐI	HREYFING	
GRAS	HEILSA	
MATUR	LITUR	
SKÓLI	LÖGUN	

Tafla 1. Dæmi um merkingarflokka í Íslenskri nútímamálsorðabók.

Flokkar eins og þeir sem sjá má í töflu 1 eru um fimm hundruð talsins og mjög misstórir. Þeir stærstu eru yfir þúsund orð en þeir minnstu aðeins nokkrir tugir orða. Hvert uppflettiorð getur tilheyrt fleiri en einum flokki. Orðið *grænkál* tilheyrir t.d. bæði flokkunum GRAS og MATUR, og *kennari* tilheyrir flokkunum SKÓLI og STARFSHEITI. Flokk-

arnir eru ekki sýnilegir notandanum en þeir voru ekki unnir á þann hátt að hentaði til birtingar. Þetta vinnulag er nýmæli í íslenskri orðabókagerð, og hefur þessi merkingarflokkun reynst mikilvægt hjálpartæki við ritstjórnarvinnuna. Meðal annars tryggir hún betra samræmi í meðferð merkingarskyldra orða þar sem hægt er að gera mismunandi blæbrigðum orða skil á markvissari hátt. Hvern flokk er hægt að skoða í innra samhengi og hugsanlega bæta í hann uppfllettiorðum til að fylla í myndina og draga úr líkum á að óæskileg göt séu í orðaforðanum. Auk þess má fá úr flokkunum margskonar aðrar upplýsingar, t.d. er hægt að búa til orðasöfn í ákveðnum greinum.⁷ Nánari fróðleik um merkingarflokkana er að finna hjá Þórdísi Úlfarsdóttur (2013:57–58).

3.2.2 Notkun textasafna

Eftir því sem fram líða stundir er orðið alsíða að setja saman orðabækur með hjálp textasafna, og á það jafnt við um Ísland og önnur Evrópulönd. Nú um stundir er raunar óhugsandi að ætla sér að gera orðabók án aðkomu textasafna. Textasafn er tiltekið magn texta á ákveðnu tungumáli og geta þeir verið af ýmsum toga. Sum textasöfn hafa að geyma fjölbreytilega textaflokka á meðan önnur eru einleitari að efni, hafa t.d. mikið umfang fréttar- eða dagblaðatexta, en textasöfn af öllu tagi koma að notum fyrir orðabókahöfunda. Á meðal nútímalegra orðabóka sem byggja mjög á textasöfnum er *Den danske ordbog* (DDO). Ritstjórar verksins hafa lagt á það áherslu að tíðni orðs geti ráðið úrslitum um það hvort orðið sé haft með í flettulistanum (sjá Jensen o.fl. 2018:143). Þá gildir einu hvort orðið er „æskilegt“ eða ekki, hvort það sé t.d. notað á niðrandi hátt um minnihlutahópa. Vinnureglan er sú að ef orðið er í mikilli notkun eiga notendur að geta fengið upplýsingar um það í DDO. Í *Íslenskri nútímamálsorðabók* er í stórum dráttum fylgt sömu stefnu.

Textasöfn með íslenskum textum hafa verið í hraðri þróun á undanförmum tveimur áratugum, en við gerð *Íslenskrar nútímamálsorðabókar* var einkum stuðst við eftirfarandi textasöfn:

- *Íslenskt textasafn* hefur að geyma um 90 milljónir lesmálsorða úr mjög margbreytilegum textaflokkum. Hægt er að leita að

⁷ Með þessari aðferð hefur þegar verið útbúið orðasafn með lögfræðihugtökum milli íslensku og frönsku (*Íslenskt-franskt lögfræðiorðasafn* 2018). Svipað er hægt að gera án mikillar fyrirhafnar með fleiri merkingarsvið orðaforðans.

öllum orðmyndum tiltekins orðs í einu en annars er þetta textasafn ekki markað málfræðilega.

- *Mörkuð íslensk málheild* (MÍM) er málfræðilega markað textasafn með 25 milljónum lesmálsorða. Textarnir eru fjölbreyttir og eru þeir allir frá árinu 2000 og yngri.
- *Tímarit.is* er stafrænt safn sem veitir aðgang að milljónum myndaðra eða skannaðra blaðsíðna úr dagblöðum og tímaritum frá upphafi vega. Blöðin hafa að geyma almennt fréttafni, auglýsingar, efni á sviði bókmennta, sagnfræði, ættfræði, þjóðlífs, menningar, atvinnuvega og viðskipta.
- *Risamálheildin* er nýjasta textasafnið fyrir íslensku og eins og nafn þess bendir til er það gríðarmikið að vöxtum, með yfir milljarð lesmálsorða. Mikið af efninu er úr fréttamiðlum, stjórnsýslu, fræðslumiðlum og bloggi (sjá Steinþór Steingrímsson, Sigrúnu Helgadóttur og Eirík Rögnvaldsson 2018).

Eins og ISLEX er *Íslensk nútímamálsorðabók* að miklu leyti grundvölluð á textasöfnum, en sú vitneskja sem sækja má þangað hefur fjölþætt gildi í orðabókavinnu:

- Textarnir gefa vísbendingu um tíðni orðsins, þ.e. hversu algengt eða sjaldgæft það er í málinu. Auk þess er í mörgum tilfellum hægt að sækja slíkar tölur beint í gagnagrunn viðkomandi textasafns.
- Upplýsingar fást um aldursdreifingu orðsins og notkunarvið þess.
- Hægt er að ákvarða merkingu orðs, eina eða fleiri, út frá samhengi þess í textanum.
- Ákveðin mynstur orða geta komið í ljós í röðuðum leitarniðurstöðum, þ.e. mynstur um fasta orðanotkun geta birst á skýran hátt.

Til eru öflug tól sem aðstoða orðabókarsmiðinn við að velja mikilvæg orð og orðasambönd úr orðaforðanum, en það gerir orðabókarvinnuna mun sjálfvirkari en áður hefur þekkt. Þetta hefur þó ekki verið notað við gerð *Íslenskrar nútímamálsorðabókar*.

4 Ritstjórnarstefna – orðaforði og orðskýringar

4.1 Orðaforði og málstefna

Þegar orðabók er búin til er mótuð ritstjórnarstefna sem fylgt er við vinnuna, bæði hvað snertir orðaforðann og orðalagið sem notað er í skýringum og notkunardæmum. Vanda þarf vel þær upplýsingar sem gefnar eru enda hafa notendur ríka tilhneigingu til að treysta því sem stendur í orðabókum.

Orðaforðinn er þungamiðja orðabókarinnar og val hans getur verið vandmeðfarið. Feitletruðu uppflettiorðin er það sem að jafnaði er mest áberandi í orðabókinni. Við upphaf vinnu við rafrænar orðabækur hjá Orðabók Háskólans (sem síðar varð hluti af SÁM) árið 2005 var nauðsynlegt að huga vel að þessum þætti. Eins og kom fram í kafla 3.1.1 var við gerð *Íslenskrar nútímamálsorðabókar* notast við sérstaklega tilbúinn orðabókarstofn sem skilgreindur var fyrir ISLEX-verkefnið, en eftir að hann hafði verið lesinn inn í gagnagrunn orðabókarinnar voru uppflettiorðin unnin áfram eins og áður segir. Í framhaldi af því var ákveðið að fara út fyrir ramma hefðarinnar hvað snerti ýmsa þætti; má þar nefna að mörg uppflettiorðin eru tvíyrt eða margyrt, ekki síst atviksliðir og samtengingar (*að fullu, þar af leiðandi*), og miðmyndir og lýsingarhættir sagna voru gerð að sérflettum. Þannig hafa orðmyndirnar *ráða, ráðast, ráðandi* og *ráðinn* allar stöðu uppflettiorðs.

Þegar hafist var handa við að gera *Íslenska nútímamálsorðabók* voru þessi atriði höfð með óbreyttum hætti. Val orðaforðans var stöðugt endurmetið en aðallega var talið nauðsynlegt að bæta inn orðum. Ákveðið var að horfa meira til nýlegra orða í tungumálinu, m.a. óformlegs orðaforða úr talmáli, orða af erlendum uppruna, svo og nýyrða. Mikið af þeim breytingum skilaði sér einnig í ISLEX-orðabókina og hina íslensk-frönsku *Lexíu*.

Það hefur löngum verið nokkur tregða við að taka inn erlend orð í íslenskar orðabækur, jafnvel þótt viðkomandi orð séu í virkri notkun hjá almenningi. Það er endurspeglun þess að í rituðu og töluðu máli hefur lengi ríkt ákveðin togstreita milli hreintungustefnu og daglegs tungutaks, eins og Íslendingar virðist almennt vera mjög meðvitaðir um (sjá t.d. umfjöllun Ara Páls Kristinssonar 2017:69 o.áfr. og 129 o.áfr.). Af þessum ástæðum þarf orðabókarhöfundur að huga að því hvenær orð af erlendum uppruna er „komið inn“ í íslensku og hefur öðlast þegnrétt í málinu. Stundum geta orðabækur raunar haft

nokkuð um það að segja þar sem þær hafa gjarnan verið ákveðnir áhrifavaldar um tungumálið.

Það má ímynda sér aðstæður árið 2018 þar sem orð af erlendum uppruna er tekið inn í orðabók vegna þess að orðið er raunverulega notað, gjarnan fyrst á óformlegum vettvangi. Menn setja stundum fyrirvara eða afsaka sig þegar þeir nota slíkt orð, og segja þá t.d. í fjölmiðlaviðtali „ef ég má sletta.“ Þróun slíks orðs í orðabók getur verið sú að orðið er kannski fyrst merkt *slangur*, síðan *óformlegt*, og ef orðið sýnir engin merki um að það sé að hverfa á braut er slíkum merkimiðum sleppt á endanum (í sumum tilfellum), enda er orðið þá orðið hluti af eðlilegri íslensku. Meðal eldri dæma um slíkt orð er *dúkka*, sem er tökuorð úr dönsku og var merkt með spurningarmerki bæði í *Íslensk-danskri orðabók* (1920–1924) og í ÍO (1983). Í fyrri bókinni táknar spurningarmerkið „erlent tökuorð sem algengt er í talmáli, einkum í bæjum; ekki viðurkennt í ritmáli.“ Í þeirri síðari táknar spurningarmerkið „vont mál, orð eða merking[u] sem forðast ber í íslensku.“ Í *Íslenskri nútímamálsorðabók* er *dúkka* ekki merkt neinu sérstöku málsniði enda mundu flestir telja það látlaust og venjulegt orð.

Dæmi um orð í *Íslenskri nútímamálsorðabók* sem hefur ekki losnað við að vera merkt *óformlegt* er nafnorðið *bús*. Það er merkt sem *vont mál* í ÍO 1983. *Bús* virðist ekki enn hafa öðlast fullan þegnrétt í málinu en það gæti stafað af merkingu þess (‘áfengi’). Elsta þekktu ritheimild um orðið *bús* er frá árinu 1950 í bókinni *Vögguvísa* eftir Elías Mar (ROH), svo að orðið á sér nokkuð langa sögu í nútímaíslensku. Raunar er upprunalega enska orðið *booze* einnig merkt óformlegt í ensku veforðabókinni *English Oxford Living Dictionaries*.

Fleiri dæmi um orð sem ljóst er að eru tökuorð (eða slettur) en heyrast samt oft (og sjást) eru *blokkera*, *formúlera*, *fókusa*, *fókusera*, *kommentera*, *púritani*, *sessjón*, *tradisjón* og *vandalismi*. Þessum orðum var bætt inn í *Íslenska nútímamálsorðabók*, einkum vegna tíðni þeirra í textasöfnum, en til að upplýsa notandann um að orðin hafi ekki hlutlausan stíl í íslensku eru þau merkt *óformleg*. Eftir því sem tímar líða fram kemur í ljós hvort málsnið þeirra breytist.

Ný orð geta átt sér allt aðrar forsendur en lýst hefur verið hér. Stundum er brýn þörf á að bæta orðum við málið vegna þess að breytilegar aðstæður, ný tækni o.s.frv. krefjast nýrra orða til að nota í stað erlendra hugtaka, orða sem falla vel að íslensku máli. Ari Páll Kristinsson (2017:149 o.áfr.) hefur nefnt þetta fyrirbæri *málauðgun*. Þessi þörf er sannarlega ekki nýtilkomin, samanber það sem áður var

haft eftir Sigfúsi Blöndal (1924:vii) um að menningarleg framþróun kallaði á „mikinn fjölda nýrra hugtaka og orða á flestum sviðum.“ Til að svara þörfinni hafa mörg ný orð verið búin til, eða þá að gömul orð hafa fengið nýja merkingu. Þekkt dæmi frá fyrri tíð eru *sími*, *tölva*, *lyfta*, *þota*, *þyrta* og *skjár*, eins og fram kemur hjá Ara Páli. Nýrri orð í málinu eru *ritrýna*, *skjáskot*, *transfólk* og *rafretta*. Þegar notaðar eru nýmyndanir af þessu tagi í töluðu eða rituðu máli er almennt lítið um að fyrirvari eða afsökunarbeiðni fylgi, líkt og þegar um afdráttarlausar erlendar slettur er að ræða þegar fólk talar í fjölmiðlum. Þessi orð hafa heldur engan merkimiða í orðabókinni þar sem ekki var talin ástæða til að gefa til kynna sérstakt málsnið með þeim.

Stundum skipta fyrirbærin um heiti, eins og t.d. *bráðamóttaka* sem áður hét *slysavarðstofa* og *leikskóli* sem nefndist áður *dagheimili*. Nöfn á ráðuneytum hafa verið breytileg milli ára (*menntamálaráðuneytið* varð *mennta-* og *menningsarmálaráðuneytið*) og hafa breytingarnar væntanlega átt sér stað vegna breytts hlutverks viðkomandi ráðuneytis, samruna tveggja eða fleiri ráðuneyta o.s.frv. Slíkar nafnabreytingar eru stjórnvaldslegar ákvarðanir og gerast skyndilega, og orðabókarhöfundar hafa stundum orðið að fylgjast með breytingum á þessum heitum til að notendur fái nýjar og réttar upplýsingar.

4.2 Samhengi íslensku, eldri orðaforði

Eins og heiti *Íslenskrar nútímamálsorðabókar* gefur til kynna miðast orðaforði hennar og notkunardæmi fyrst og fremst við þann tíma sem hún er unnin á. Samt var talið nauðsynlegt að einskorða efnið ekki alveg við nútímaíslensku þar sem tungumálið stendur í raun ekki einangrað í samtíð sinni, og nútímaíslenska verður ekki svo auðveldlega aðskilin frá máli fyrri áratuga og alda. Því er einnig að finna eldri orðaforða í orðabókinni. Meira er fjallað um þennan þátt í íslenskum orðabókum hjá Þórdísi Úlfarsdóttur og Kristínu Bjarnadóttur (2017:4).

Í orðabókinni eru mörg algeng orð úr fornu máli og skáldamáli (*dreyri*, *drösull*, *fold* og *mögur*), og enn fremur talsverður orðaforði frá 20. öld sem er ekki í mikilli notkun nú á dögum (*framreiðslustúlka*, *stælgæi*, *sóttkveikja* og *stöðvarbill*). Einnig koma fyrir orð sem tilheyra gamla bænda- og fiskveiðisamfélaginu, t.d. *fráfærur*, *búandkarl*, *fjósbaðstofa*, *klyfberi*, *tómthúsmaður*, *fiskhjallur* og *útræði*. Þarna er m.a. leitast við að koma til móts við þarfir skólanemenda, sem eru mikilvægur markhópur verksins, og annarra þeirra sem vilja geta flett upp eldri

orðaförða, t.d. eins og hann birtist í verkum þekktra rithöfunda. Með því að gefa gaum að samhengi tungumálsins eykst notagildi orðabókarinnar þrátt fyrir að hún sé kennd við nútímamálið.

4.3 Nöfn í orðabókinni

Orðabækur fara mjög mismunandi leiðir hvað snertir fjölda og flokka nafna í uppflöttiorðaförðanum. Í *Íslenskri nútímamálsorðabók* eru nöfnin þau sömu og í ISLEX-orðabókinni en þar var talið mikilvægt að taka með nöfn sem eiga sér breytileg heiti milli íslensku og annarra norðurlandamála. En nöfn í einsmálsorðabók geta líka veitt gagnlegar upplýsingar, t.d. um staðarheiti sem eiga sér sérstakt nafn á íslensku. Í sumum tilfellum eru þau upprunnin í fornu máli, eins og *Hróarskelda* (Roskilde), *Óðinsvé* (Odense), *Brándheimur* (Trondheim) og *Kænugarður* (Kiev). Einnig er sterk hefð fyrir því að aðlaga ákveðin borgarheiti að íslensku eða stafsetja þau með íslenskum hætti: *Barselóna*, *Berlín*, *París*, *Róm*.

Fjöldi sérnafna er nánast endalaus, bæði örnefni, mannanöfn, heiti á fyrirtækjum, stofnunum og fleiri flokkar nafna. Of mikið af slíku efni gæti nánast kaffært venjulegan orðaförða í orðabók. Því hefur verið reynt að velja vandlega slíkan orðaförða og stilla fjölda nafnanna í hóf. Mannanöfnum er sleppt alveg, sem er gagnstætt því sem gert er í ÍO 2007 og fleiri útgáfum þess rits. Þeir notendur sem hafa sérstakan áhuga á mannanöfnum eða örnefnum eru oft betur settir með að fletta þeim upp í sérhæfðari ritum (sjá umfjöllun um nöfn í almennum orðabókum hjá Þórdísi Úlfarsdóttur 2017).

4.4 Sérmerkingar og málsnið

Það er viðtekin venja í orðabókum að merkja sérstaklega orð sem tilheyra afmörkuðum hluta orðaförðans og er *Íslensk nútímamálsorðabók* þar engin undantekning. Um er að ræða tvenns konar sérmerkingar sem eru af ólíku tagi: annars vegar merkingar á sértækum fagorðaförða og ídorðum í ýmsum greinum, hins vegar merkingar á málsniði. Notkun á þessum sérmerkingum er sem hér segir.

Sértækur orðaförði og fræðiorð eru venjulega merkt því fagi eða fræðigreini sem þau tilheyra, t.d. lögfræði, eðlisfræði og bókmenntafræði. Þannig er t.d. *algebra* merkt stærðfræði og *beinhimna* er merkt líffræði/læknisfræði. Tuttugu og fimm mismunandi fög eru merkt á þennan hátt:

bókmenntafræði, bragfræði, dýrafræði, eðlisfræði, efnafræði, goðafræði, grasafraði, handritafræði, heimspeki, iðnaður, jarðfræði, líffræði, læknisfræði, lögfræði, málfræði, myndlist, stærðfræði, tölvur, sögulegt, tónlist, tölvur, veðurfræði, viðskipti/hagfræði, vistfræði og þjóðtrú.

Alls eru 2.900 orð (eða merkingarliðir innan flettugreinar) merkt sérstöku fagi eða tæp 6% orðaforðans. Því má bæta við að heiti á plöntum og dýrum, fuglum og fiskum hafa yfirleitt ekki sérmerkingu enda tilheyra þau í flestum tilvikum almennum íslenskum orðaforða.

Hvað snertir málsnið var áður (í kafla 4.1) minnst á það í tengslum við orð sem eru merkt *óformleg* í flettuorðaforðanum. Stundum er talin ástæða til að upplýsa notandann ef orð tilheyrir ákveðnu málsniði, þ.e. orðið hefur ekki hlutlausan stíl. Orðið kann að vera gildishlaðið á neikvæðan hátt, t.d. *fullirafstur*, *greppitrýni* og *senditík*. Í slíkum tilvikum er orðið merkt niðrandi. Sum orðin eru merkt gamaldags, t.d. *blankskór*, *býtta* og *forkelast*. Önnur orð eru talin formleg: *baðmull*, *bifreið*, *fóthemill*. Tuttugu og fjórar mismunandi málsniðsmerkingar eru notaðar í orðabókinni:

barnamál, formlegt, fornt, gamaldags, gamalt, gamansamt, gróft, gæluorð, háð, hátíðlegt, í ávarpi, í eldra máli, niðrandi, oft í háði, óformlegt, ekki fullviðurkennt mál, niðrandi, óvirðulegt, sjaldgæft, skáldamál, staðbundið, til áherslu, úrelt og yfirfærð merking.

Alls eru um 1.800 orð merkt sérstöku málsniði. Þeim sem áhuga hafa er bent á almennan fróðleik um málsnið íslensku hjá Ara Páli Kristinssyni (2017:188–191).

4.5 Orðskýringar

Orðskýringar eru grundvallarþáttur einsmálsorðabókar og þær þurfa að vera vandaðar. Í *Íslenskri nútímamálsorðabók* eru 62.400 orðskýringar. Uppflettiorðin eru 50.000 og stafar munurinn af því að sum orðin hafa fleiri en eina merkingu, auk þess sem skýringar eru við orðasamböndin. Þau orð sem hafa flesta merkingarliði, og þar með flestar skýringar, eru sagnirnar *taka*, *koma* og *fara*, en hver þeirra hefur marga tugi merkingarliða. Þau nafnorð sem hafa flesta merkingarliði eru *brot*, *efni*, *hlaup* og *mark*.

Orðskýringarnar eiga rót sína að rekja til ISLEX-verkefnisins. Þar

höfðu átta manns unnið að skýringunum og var því fremur ólíkur stíll á þeim eftir því hver hafði samið þær. Þær voru einnig mjög misvandaðar, enda fyrst og fremst hugsaðar til hjálpar þýðendunum. Þær gegndu engu öðru hlutverki í þeirri orðabók og voru ósýnilegar notendunum. Samt voru settar skýringar við öll uppflöttiorðin, ein eða fleiri eftir fjölda merkingarliða viðkomandi orðs, og það kom í ljós að sumar þeirra voru ágætlega nothæfar í *Íslenska nútímamálsorðabók*. En til þess að hægt væri að birta skýringarnar þurfti fyrst að skoða þær með gagnrýnum augum, og var víða nauðsynlegt að endurskoða og samræma skýringarnar og breyta orðalagi þeirra.

Í *Íslenskri nútímamálsorðabók* er gert ráð fyrir að notandinn þekki grunnorðaforða íslensku, því að einhvern veginn þarf að vera hægt að orða skýringarnar. Ritstjórar orðabókarinnar hafa haft að leiðarljósi markhópa eins og t.d. nemendur í íslenskum skólum og þá sem hafa lært íslensku sem annað mál. Þegar orðskýringar eru samdar þarf að hafa í huga nokkrar grundvallarreglur. Það er venja í orðabókum að hafa orðalag skýringa eins skýrt og látlaust og kostur er, og laust við sjaldgæf orð. Ekki er felldur dómur í skýringunni, t.d. er ekki notað orðalag eins og 'mjög fallegur mófugl' eða 'viðbjóðslegur glæpamaður'. Skýringin þarf að vera stutt en samt nógu löng til að notandinn geti fræðst um hvað orðið merkir. Lorentzen og Trap-Jensen (2012:96) hafa bent á að vandasamt sé að finna jafnvægi milli þess að hafa stutta og gagnorða skýringu við tiltekið orð, og þess að veita nægar upplýsingar. Þeir nefna að skólanemendum sé best þjónað með einföldum skýringum fremur en að þær séu langar og nákvæmar.

Orðabókarskýring þarf að geta rúmast í einni setningu og ekki er hafður punktur í lokin. Kommur og aukasetningar eru þó leyfðar í skýringunni. Að þessu leyti er reginmunur á framsetningu upplýsinga sem hefðbundnar orðabækur veita og hins vegar alfræðirit. Í síðar-nefndu ritunum eru ekki sömu hömlur á lengd skýringa og í orðabókum, og þar er flæði textans mun frjálsara og stíllinn líkari því sem almennt gerist um samfelld mál.

Loks má minnst á orðasambönd. Mikilvægt er að skýra orðasamböndin þar sem merking þeirra er oft algerlega ógegnsæ og veita þarf notandanum upplýsingar um hvað þau þýða. Oft eru orðasambönd fremur erfiður þáttur tungumáls varðandi málskilning, má t.d. nefna orðasambönd eins og *leggja upp laupana*, *bera ekki sitt barr*, *fara ekki í grafgötur um e-ð*, *hafa ekki roð við e-m* og *geta trútt um talað*. Sum orðasambönd eru ekki algeng í málinu og eru jafnvel að hverfa úr virkri

málnotkun, en engu að síður bregður þeim fyrir og er því mikilvægt að þau séu finnanleg í orðabókum.

4.6 Notkun og uppfærslur

4.6.1 Notkun orðabókarinnar

Íslensk nútímamálsorðabók hefur sína eigin vefsíðu (*islenskordabok.is*), og hún er jafnframt eitt af þeim gagnasöfnum sem liggja að baki síðunni *málið.is*. Sá vefur var opnaður haustið 2016 og var orðabókin tekin í notkun við sama tækifæri, þótt hún væri ekki fullgerð. Vefgáttinni er ætlað að fræða og leiðbeina um íslenskt mál öllum þeim sem þurfa á því að halda (sjá meira um það hjá Steinþóri Steingrímssyni, Ara Páli Kristinssyni og Halldóru Jónsdóttur 2018), og hafa notendur tekið vel þessu nýja hjálpartæki. Margir af notendum nútímamálsorðabókarinnar hafa uppgötvað hana með því að fara inn á *málið.is*. Á þeim tíma sem *Íslensk nútímamálsorðabók* hefur verið aðgengileg hefur notendum hennar fjölgað jafnt og þétt, og voru þeir tæplega 130.000 árið 2018 samkvæmt talningu. Notendur eiga þess kost að koma með spurningar og athugasemdir um orðabókina í sérstökum reit á heimasíðunni, og hafa sumar þeirra leitt til viðbóta og leiðréttinga á efninu.

4.6.2 Uppfærslur og framhald

Eftir að veforðabækur urðu allsráðandi form orðabóka upp úr aldamótunum 2000 hefur dregið mjög úr áhuga bókaforlaga á að gefa út prentaðar orðabækur, enda hefur sala þeirra dregist verulega saman. Veforðabækur hafa nánast tekið völdin og margar þeirra hafa ókeypiss aðgang. En ef orðabók um nútímamálið á vera gott hjálpargagn og raunsönn heimild um tungumál samtímans þarf að vera vilji og fjármagn fyrir hendi til að viðhalda verkinu því að uppfærslur kosta mikla vinnu. Þróun og viðhald orðabóka á rafrænu formi er að sjálfsögðu talsvert einfaldara mál heldur en þegar gefin er út prentuð bók. Það táknar þó ekki að slíkt gerist nánast af sjálfu sér, þar sem beytingar á efninu útheimta að virk ritstjórn starfi við orðabókina sem annast lagfæringar og viðbætur.

Í upphafi var ætlunin að halda stærð *Íslenskrar nútímamálsorðabókar* í kringum 50.000 orð en hugmyndir eru nú uppi um stækkun hennar til að koma betur til móts við þarfir notenda. Ekki er enn vitað hver

sú þróun verður, en það er ljóst að mikil notkun orðabókarinnar gefur til kynna að brýn þörf er á að uppflettirit af þessu tagi sé til staðar, reglulega uppfært og öllum aðgengilegt.

5 Lokaorð

Eins og komið hefur fram hér að framan er *Íslensk nútímamálsorðabók* ný orðabók, ætluð nýjum tímum og þeim notendum sem leita sér upplýsinga um tungumálið á vefnum. Orðabókin byggist að verulegu leyti á margmála veforðabókinni ISLEX og er upphaflegi orðalistinn fenginn þaðan ásamt mörgum öðrum þáttum í því verki. Við gerð orðabókarinnar var horft til þess að tungumálið er í stöðugri þróun og tekur breytingum, og hefur verið leitast við að láta orðaforðann og notkunardæmin taka mið af því hvernig íslenska er raunverulega töluð og rituð af málnotendum samtímans.

Nýjungar við gerð *Íslenskrar nútímamálsorðabókar* eru einkum fólgnar í þremur meginþáttum, sem hver um sig hefur afgerandi þýðingu fyrir bæði vinnslu og endanlega útkomu efnisins. Þessir þættir eru í fyrsta lagi vinnsla flettugreina eftir merkingarflokkum, í öðru lagi notkun textasafna, og í þriðja lagi hönnun fyrir birtingu á vefnum.

Vinnsla flettugreina eftir merkingarflokkum er tiltölulega ný aðferð við orðabókavinnu en hún hefur í för með sér margvíslegan ávinning. Nefna má að aðferðin skilar betra samræmi milli skyldra og hliðstæðra orða auk þess sem hún getur komið í veg fyrir óæskileg göt í orðaforðanum.

Textasöfn hafa skipt höfuðmáli við uppbyggingu *Íslenskrar nútímamálsorðabókar* og er hún fyrsta íslenska einsmálsorðabókin sem er að verulegu leyti samin með hliðsjón af þeim upplýsingum sem fást úr textasöfnum. Þar má nefna tíðni orða í málinu, merkingu þeirra, stöðu í setningum og föst orðasambönd.

Orðabókin var frá upphafi hönnuð fyrir rafræna birtingu og hefur framsetning efnisins tekið mið af því. Þess hefur verið gætt að láta kosti vefbirtingar njóta sín og kemur það m.a. fram í framsetningu efnisins, notkun margmiðlunarefnis, s.s. hljóðskráa, upplýsingum um beygingar með tenglum í annan gagnagrunn, fjölda skýringarmynda ásamt ítarleit í texta orðabókarinnar. Ofangreind vinnubrögð hafa haft í för með sér að öll orðabókavinnan er fljótlegri og auðveldari en áður var, til dæmis er ekki lengur þörf á að lesa heilu fræðibækurnar, stjórnsýslutexta og önnur rit til þess að fanga orðaforðann og merkingar orðanna.

Að lokum má nefna að það hefur sýnt sig að ýmislegt efni úr *Íslenskri nútímamálsorðabók* getur nýst til ýmissa máltæknilegra verkefna.⁸ Gera má ráð fyrir að orðabókin haldi áfram að vaxa og dafna en grundvöllurinn fyrir því er þó ætíð háður ytri aðstæðum.

Heimildir

- Ari Páll Kristinsson. 2017. *Málheimar: sitthvað um málstefnu og málnotkun*. Reykjavík: Háskólaútgáfan.
- BÍN = Beygingarlýsing íslensks nútímamáls. Kristín Bjarnadóttir (ritstj.). *Stofnun Árna Magnússonar í íslenskum fræðum*. <http://bin.arnastofnun.is> (ágúst 2018).
- Björn Halldórsson. 1992 [1814]. *Orðabók: íslensk – latnesk – dönsk*. Eftir handriti í Stofnun Árna Magnússonar í Kaupmannahöfn. Jón Aðalsteinn Jónsson sá um útgáfuna. Orðfræðirit fyrri alda II. Reykjavík: Orðabók Háskólans.
- DDO = *Den danske ordbog*. Det danske sprog- og litteraturselskab. <https://ordnet.dk/ddo> (ágúst 2018).
- English Oxford Living Dictionaries*. Oxford: Oxford University Press. <https://en.oxforddictionaries.com/> (ágúst 2018).
- Guðmundur Andrésón. 1999 [1683]. *Lexicon Islandicum*. Ný útgáfa. Gunnlaugur Ingólfsson og Jakob Benediktsson önnuðust útgáfuna. Orðfræðirit fyrri alda IV. Reykjavík: Orðabók Háskólans.
- Halldóra Jónsdóttir og Þórdís Úlfarsdóttir. 2016. Hvem bruger ISLEX og hvordan? *LexicoNordica* 23:89–104.
- ISLEX-orðabókin*. 2011. Þórdís Úlfarsdóttir (ritstj.). Stofnun Árna Magnússonar í íslenskum fræðum. <http://islex.is/> (ágúst 2018).
- ÍO 1963 = *Íslensk orðabók handa skólum og almenníngi*. 1963. Árni Böðvarsson (ritstj.). Reykjavík: Menningarsjóður.
- ÍO 1983 = *Íslensk orðabók handa skólum og almenníngi*. 1983. Árni Böðvarsson (ritstj.). Reykjavík: Menningarsjóður.
- ÍO 2002 = *Íslensk orðabók*. 2002. (3. útg. aukin og endurbætt). Mörður Árnason (ritstj.). Reykjavík: Edda.
- ÍO 2007 = *Íslensk orðabók*. 2007. (4. útg.). Mörður Árnason (ritstj.). Reykjavík: Edda.
- ÍO = *Íslensk orðabók* (á vefnum). Mörður Árnason (ritstj.). Reykjavík, Forlagið. <https://snara.is/> (ágúst 2018).
- Íslensk nútímamálsorðabók*. Halldóra Jónsdóttir og Þórdís Úlfarsdóttir (ritstj.). Stofnun Árna Magnússonar í íslenskum fræðum. <http://islenkordabok.is/> (ágúst 2018).
- Íslenskt-franskt lögfræðiorðasafn*. 2018. Valin orð úr Lexíu, íslensk-franskri vef-orðabók. Þórdís Úlfarsdóttir (ritstj.). Stofnun Árna Magnússonar í íslensk-

⁸ Slíkt verkefni er í vinnslu hjá Kristínu Bjarnadóttur og Kristínu I. Hlynsdóttur (2018–2019) sem nefnist BÍN-kjarninn.

- um fræðum. <http://lexia.arnastofnun.is/d/logfraedi2.pdf> (ágúst 2018).
- Íslenskt textasafn. Þórdís Úlfarsdóttir (ritstj.). *Stofnun Árna Magnússonar í íslenskum fræðum. corpus.arnastofnun.is/* (ágúst 2018).
- Jensen, Jonas, Henrik Lorentzen o.fl. 2018. *Thaipiger, muskelhunde og fulde svenskere: nedsættende ord, stereotyper og ligestilling i den danske ordbog. Í: Nordiske Studier i Leksikografi. Rapport fra 14. Konference om Leksikografi i Norden, Reykjavík 30. maj – 2. juni 2017:141–151.* Reykjavík: Skrifter udgivet af Nordisk Forening for Leksikografi.
- Jón Árnason. 1994 [1738]. *Nucleus latinitatis.* Guðrún Kvaran og Friðrik Magnússon önnuðust útgáfuna. Orðfræðirit fyrir alda III. Reykjavík: Orðabók Háskólans.
- Jón Hilmar Jónsson. 2005. *Stóra orðabókin um íslenska málnotkun.* Reykjavík: JPV.
- Jónas Jónasson. 1896. *Ný dönsk orðabók með íslenzkum þýðingum.* Reykjavík: Ísafoldarprentsmiðja.
- Konráð Gíslason. 1851. *Dönsk orðabók með íslenzkum þýðingum.* Kaupmannahöfn.
- Kristín Bjarnadóttir. 1998. *Norræna verkefnið.* <https://notendur.hi.is/~kristinb/#rit> (ágúst 2018).
- Lexía. *Íslensk-frönsk veforðabók (í vinnslu).* Þórdís Úlfarsdóttir (ritstj.). Stofnun Árna Magnússonar í íslenskum fræðum. <http://lexia.arnastofnun.is/> (ágúst 2018).
- Lorentzen, Henrik og Lars Trap-Jensen. 2012. Nödvendigt, tilstrækkeligt, typisk? – nogle tanker om definitionspraksis. *LexicoNordica* 19: 79–101.
- MÍM = *Mörkuð íslensk málheild.* Sigrún Helgadóttir (ritstj.). Stofnun Árna Magnússonar í íslenskum fræðum. <http://mim.arnastofnun.is/> (ágúst 2018).
- Nordisk leksikografisk ordbok.* 1997. Bergenholtz, Henning o.fl. (ritstjórar). Osló: Universitetsforlaget AS.
- Ordbog over det Danske Sprog.* 1.–28. bindi. 1919–1956. Kaupmannahöfn: Gyldendalske boghandel, Nordisk forlag.
- Ordbok över svenska språket,* utgiven av Svenska Akademien. 1898–. Lund.
- Oxford English Dictionary.* 1884–1928 (1. útgáfa). Oxford: Clarendon Press.
- Risamálheildin.* Stofnun Árna Magnússonar í íslenskum fræðum. <http://malheildir.arnastofnun.is/> (ágúst 2018).
- ROH = *Ritmálssafn Orðabókar Háskólans.* Stofnun Árna Magnússonar í íslenskum fræðum www.arnastofnun.is/page/arnastofnun_gagnasafn_ritmal (ágúst 2018).
- Sigfús Blöndal. 1920–1924. *Íslensk-dönsk orðabók.* Reykjavík: Íslensk-danskur orðabókarsjóður.
- Sigfús Blöndal. 1963. *Íslensk-dönsk orðabók. Viðbætur.* Ritstjórar: Halldór Halldórsson og Jakob Benediktsson. Reykjavík: Íslensk-danskur orðabókarsjóður.

- Steinþór Steingrímsson, Sigrún Helgadóttir og Eiríkur Rögnvaldsson. 2018. An Icelandic Gigaword Corpus. *Nordiske Studier i Leksikografi*. Rapport fra 14. Konference om Leksikografi i Norden, Reykjavík 30. maj – 2. juni 2018:246–254. Reykjavík: Skrifter udgivet af Nordisk Forening for Leksikografi.
- Steinþór Steingrímsson, Ari Páll Kristinsson og Halldóra Jónsdóttir. 2018. Málið.is. *Nordiske Studier i Leksikografi*. Rapport fra 14. Konference om Leksikografi i Norden, Reykjavík 30. maj – 2. juni 2018:255–261.
- Þórdís Úlfarsdóttir. 2013. ISLEX – norræn margmála orðabók. *Orð og tunga* 15. Reykjavík: Stofnun Árna Magnússonar í íslenskum fræðum.
- Þórdís Úlfarsdóttir. 2017. Proprier som opslagsord – kriterier for almenordbøger. *LexicoNordica* 24:131–149.
- Þórdís Úlfarsdóttir og Kristín Bjarnadóttir. 2017. The lexicography of Icelandic. Í Patrick Hanks og Gilles-Maurice De Schryver (ritstj.). *International Handbook of Modern Lexis and Lexicography*, bls. 1–3. Berlin, Heidelberg: Springer.
- Timarit.is*. Tímaritavefur Landsbókasafns Íslands – Háskólabókasafns. <http://timarit.is> (ágúst 2018).

Lykilorð

orðabókafræði, íslenska, veforðabók, orðskýringar, merkingarsvið

Keywords

lexicography, Icelandic, online dictionary, definitions, semantic fields

Abstract

Íslensk nútímamálsorðabók (*A Dictionary of Contemporary Icelandic - islenskordabok.is*) is a new dictionary designed for the web. It is largely modelled on the multilingual ISLEX project (www.islex.is), inheriting its lemmalist as well as its general structure, such as division into meanings, examples of use, information on the valency of verbs and prepositions, the use of labels, pronunciations, illustrations etc. The fact that a living language is always open to change has affected the work, so the dictionary's lemma list and other contents reflect the spoken and written Icelandic of today, this also entails regular updates and additions to the material. The article first discusses the predecessors of *Íslensk nútímamálsorðabók*, it also briefly deals with the publishing history of Icelandic dictionaries. A description is given of the dictionary's framework, linguistic approach, contents and its main features, as well as the editorial policy of the work as reflected in its vocabulary and definitions, and the way it deals with loan words and slang. Three important aspects of the dictionary come into focus which all can be said to be groundbreaking in Icelandic lexicography. These are: the processing of all articles according to the lemma's semantic fields, a heavy use of corpora, and the dictionary's original design for the web. Finally, some information is given on the use of the dictionary since its opening in late 2016.

*Halldóra Jónsdóttir
Stofnun Árna Magnússonar í íslenskum fræðum
Laugavegi 13
IS-101 Reykjavík
halldora.jonsdottir@arnastofnun.is*

*Þórdís Úlfarsdóttir
Stofnun Árna Magnússonar í íslenskum fræðum
Laugavegi 13
IS-101 Reykjavík
thordis.ulfarsdottir@arnastofnun.is*

